

All files and numbers correspond to Figure 3.

1 - Transcript abundance table: The table of normalized transcript counts must have one row for each transcript, and one column for each individual replicate in the experiment. Row names (transcript IDs) must be unique, and column names (replicate IDs) must be unique and match the experimental design metadata (Files #2 and #5).

2 - Experimental design metadata table: The experimental design table must include a row for each replicate, and must include the following columns: “replicate.id”, “treatment.id”, and “treatment.name”. Replicate IDs must match the column names of the transcript abundance table (File #1). Treatment names should be meaningful to a new reader exploring the DrEdGE website. Treatment IDs can be the same as treatment names, or can be codes to keep file names shorter. No row names are necessary.

3 - Directory of all pairwise comparison data tables: The directory of pairwise comparison tables must contain a unique file for each unique pair of treatments. Each file must have a table with a row for each transcript, and the following three columns in this order: (1) the logarithm of the fold change in transcript abundance between the first and second treatments, (2) the logarithm of the average transcript abundance amongst all replicates of both treatments, and (3) the P-value for fold change described in column 1. Row names (transcript IDs) must match the row names in the transcript abundance file (File #1). The names of the files must reference the treatment IDs used in the experimental design metadata (Files #2 and #5). The author is encouraged to use his or her own preferred analytic methods to create the pairwise comparison tables, but he or she is also welcome to run the `pairwise_comparisons.R` script provided in the DrEdGE package. This script receives the transcript abundance table (File #1) and the experimental design file (File #2) as input, and uses edgeR to generate, name, and export all possible pairwise comparisons. The edgeR analysis models the data as a negative binomial distribution, and models overdispersion as a Poisson model. An empirical Bayes procedure is used to moderate the degree of overdispersion across transcripts, and differential expression is assessed using an exact test adapted for overdispersed data²⁸.

4 - Minima and maxima from pairwise comparison tables: These minima and maxima will be used to set the axis limits for the MA plot on the DrEdGE website. The numbers will be entered manually, so format is not important. If the author has run the `pairwise_comparisons.R` script to generate the directory of pairwise comparison tables (Files #3), these minima and maxima values will automatically be calculated and a report will be generated.

5 - Experimental design metadata JSON: The experimental design metadata must also be provided as a JSON (JavaScript Object Notation) file in which each first-order object describes each treatment from the experiment. The `JSON_treatments.R` script provided in the DrEdGE package will accept File #2 and generate this JSON. If the author wishes to generate their own JSON file, the first order object keys must be the short unique codes for the treatment (“treatment.id” in File #2), and the values must include the following second order objects: (1) A “label” key that points to a verbose name for the treatment (“treatment.name” in File #2), and a “replicates” key that points to the full list of replicate IDs for the treatment (“replicate.id” in File #2, which must each match a column name from the transcript abundance table).

6 - Transcript synonyms or alternate names: The transcript names that are optimal for the transcript abundance table may be coded, and unevocative to a reader exploring the DrEdGE website. The author may wish to associate a preferred “readable” name with each transcript, and/or include other code names that readers might use in the search bar. For this reason, we have provided the option to include a file of alternative names, or synonyms, for transcripts. With one line of the file per transcript, each line must begin with the transcript name that the author prefers the DrEdGE visualization to show. Following that first-choice name, the author should list all other names or codes for the transcript, separated by tabs. In the input files for our three example websites (Supplementary File 1), we have included transcript synonym tables for human, mouse, and worm.

7 - Organism-specific database search URL: If the author would like transcripts to link to an external database (i.e. Wormbase³⁷), he or she may provide a generic search URL for this database, with “%name” inserted where the transcript name would be. For example, the generic search URL for Wormbase would be <http://wormbase.org/search/gene/%name>.

8 - Custom grid layout for grid heat map: The default settings represent each treatment as a box icon, with boxes fit together in a compact grid (Figure 2A). Alternatively, the author can specify how he or she would like these boxes arranged (Figure 2B,D). To do this, the author must create a comma delineated file in which each row of text describes the sequence of treatments (or white space) for each row of the grid.

9 - Custom SVG for illustrated heat map: If the author wishes to display the heatmap as a graphic illustration, he or she may upload an SVG (Scalable Vector Graphics) file with custom icons representing each sample (Figure 2C,E). The author must make create a vector file with one object for each treatment, and each object must be given a name that matches the treatment ID used in the experimental design metadata files (Files #2 and #5).